When Expectation-maximization-based Theories Work or Do Not Work: An Eye-Tracking Study of the Discrepancy between Everyone and Every One*

LIU Hong-Zhi[3] LI Xingshan[1, 2] LI Shu[1, 2, 4] RAO Li-Lin[1, 2]

([1]CAS Key Laboratory of Behavioral Science, Institute of Psychology, Chinese Academy of Sciences, Beijing 100101, China)

([2]University of Chinese Academy of Sciences, Beijing 100049, China)

([3]Department of Social Psychology, Zhou Enlai School of Government, Nankai University, Tianjin 300350, China)

([4]Department of Psychology and Behavioral Sciences, Zhejiang University, Hangzhou 310028, China)

Abstract Mainstream theorists in risky decision-making have developed various expectation theories with the ambitious goal of capturing everyone's choices. However, ample evidence has revealed that these expectation theories could not capture every individual's ("every one's") actual risky choice as descriptive theories. With doubts about the default compatibility between everyone (full set) and every one (subset), we used an eye-tracking technique to explore whether a theory for everyone would work well for every one. We found that expectation theories could capture the choice of an individual when making decisions for everyone and for self in a multiple-play condition, but could not capture the choice of an individual when making decisions for self in a single-play condition. Our findings contribute to a better understanding of the boundaries of expectation theories and those of heuristic/non-expectation models, and may shed light on the general issue of the classification of risky decision-making theories.

Keywords risky choice, decision for everyone, expectation-maximization, discrepancy between everyone and every one, eye-tracking

---

* Received: 2022-01-22

Supported by: National Natural Science Foundation of China (71901126); National Social Science Foundation of China (19ZDA358); Humanity and Social Science Youth Foundation of Ministry of Education of China (19YJC190013); Fundamental Research Funds for the Central Universities (63222045).

Corresponding author: RAO Li-Lin, E-mail: raoll@psych.ac.cn

1 Introduction

In an apocryphal story, Howard Raiffa, an influential Bayesian decision theorist and pioneer in decision analysis, was on the faculty at Columbia in 1957 when he received an offer from Harvard. He visited a friend, the dean at Columbia, and asked for advice. Sarcastically, the dean replied that Raiffa should identify the relevant criteria, weight each criterion, rate each school on each criterion, do the arithmetic, see which school had the best overall score, and go there. Raiffa's response was, "No, this is a serious decision." (See Bazerman et al., 1998). This story presents the potential for a discrepancy between a type of decision theory that is developed to describe everyone's choice and a type of decision theory that is developed to describe an individual's choice. That is, mainstream decision theorists (especially economic theorists) have developed various "weighting, adding, and maximizing" theories with the ambitious goal of capturing everyone's choice, but, when faced with a choice in the real world, decision-makers (even theorists themselves) are unwilling to apply those sophisticated weighting and adding theories to their own choices. Inspired by Raiffa's story, the goal of the present study was to explore the mechanism underlying *the discrepancy between a theory for everyone and a theory for self in decision-making under risk.*

In the field of decision-making under risk, the history of expectation theories, which involve maximizing an expectation, very much reflects Raiffa's story. The mainstream theories of decision-making under risk have never abandoned the core framework of expectation, which assumes a weighting and adding process (Payne & Braunstein, 1978). According to expectation theories, for each option, individuals weigh the utility of each outcome by some function of probability, sum up all weighted utilities, and select the option that offers the highest sum of weighted utilities (Kahneman & Tversky, 1979; Payne & Braunstein, 1978; Tversky & Kahneman, 1992). In this sense, the expectation difference between options would influence the cognitive process required by weighting and adding theories (i.e., expectation theories), but would not necessarily influence a non-weighting and adding process. The concept behind expectation theories is that, if the expectation difference between two options is relatively large, individuals can easily identify the option with the higher expectation; however, if the expectation difference between two options is small enough, individuals should find it difficult to identify the option with the higher expectation and thus experience decision conflict (Rao et al., 2011).

However, ample evidence has revealed that the expectation theories developed by mainstream theorists cannot capture an individual's (even an individual like Raiffa) actual decision under risk. Substantial research has demonstrated that people do not follow the logical process suggested by expectation theories when making risky choices, but rather rely on simplifying heuristics (Brandstätter et al., 2006; Li, 2004, 2016; Huang et al., 2021), which are guided by a distinctly different theoretical orientation. By following a heuristic/non-expectation process, people do not need to integrate information from all dimensions to arrive at a decision. Many heuristic theories, such as the maximax heuristic (Savage, 1951), the equate-to-differentiate model (Li, 2004, 2016) and the priority heuristic (Brandstätter et al., 2006), assume that people identify the differences between options in each dimension, compare the differences, and make decisions relying on only one key dimension. According to the heuristic/non-expectation theories, a change in the magnitude of the difference in the key dimension between a pair of options would be expected to result in a change in the intra-dimensional evaluation process, as assumed by the heuristic/non-expectation strategy. However, such a change would not influence the process of weighting and adding all of the possible payoffs to assign an expectation, as assumed by expectation theories.

From our point of view, the reason why mainstream decision theorists did not abandon the framework of expectation might be that these theorists never doubted the validity of the expectation rule as a descriptive rule in describing decision-making under risk (Baron, 1986). It is our belief that expectation theories may capture risky choices when individuals make decisions for everyone (decision-for-everyone) or in multiple-play conditions (decision-for-self-multiple-play), but whether these theories could capture risky choices when individuals make decisions for themselves (decision-for-self-single-play) cannot be taken for granted.

Expectation theories were developed for everyone and multiple-play conditions, under the guidance of the law of large numbers, which refers to the fact that the frequency of a random event is approximately equal to the probability of the event's occurrence after repeated decisions or in large samples (van der Stoep & Seifert, 1994). In research about risky choice, two different kinds of large numbers can be identified. One is multiple repeated realizations. If someone plays a choice game a sufficiently large number of times, the outcome will fall close to the long-run expected value (Klos et al., 2005). The other case is making decisions for a large number of people. If each decision-maker faced the same risky choice and made the same risky decision, the risk could disappear upon

aggregation and tend toward the overall expected value. In other words, multiple-play choices and decisions for a large number of people meet the unlimited repetition requirement of theoretical expectation theories. Therefore, decision-makers are prone to treating possible outcomes as fungible, taking both the statistics of the past and the multiple opportunities of the future into account. In contrast, decision-for-self-single-play could hardly meet the unlimited repetition requirement of theoretical expectation theories. Therefore, decision-makers are prone to treating problems as unique (Kahneman & Lovallo, 1993), perceive possible outcomes as nonfungible (DeKay et al., 2006; DeKay & Kim, 2005), and wisely neglect the expectation (Kahneman & Lovallo, 1993). In this case, individuals would not adopt an EV or expectation strategy.

Our speculation has been partially supported by existing behavioral and neurological evidence, which has consistently reported that the EV strategy performs adequately when people make risky choices for themselves in a multiple-play condition (Klos et al., 2005; Li, 2003; Rao et al., 2013; Su et al., 2013; Sun et al., 2013). However, surprisingly little work has been done to examine whether the expectation strategy performs adequately when making risky choices for a large number of people (everyone). Importantly, decisions for a large number of people (e.g., the government approving a new policy for everyone in the nation) might be more influential than a decision for oneself in daily life. In addition, examining the difference between decisions for everyone and decisions for self also help us understand the underlying mechanisms of some decision conflicts occur in reality. Taking medical decision-making as an example, doctors may use expectation strategies to reduce “decision noise” when making decisions for many patients (Kahneman, et al., 2021), while they may adopt heuristic strategies when making medical decisions for themselves or their family members (Popovic , et al., 2019). This discrepancy due to different decision goals or situations may lead to conflict in the doctor-patient relationship. Therefore, examining the characteristics and underlying mechanisms of decisions for everyone has both theoretical significance and practical application value.

The purpose of the present study was to use a within-subject design to examine whether the decision strategy that individuals adopt when making decisions for everyone (a large number of people) or for the multiple-play condition is the same as the strategy they use when making decisions for themselves. We assumed that, even for the same individual, he/she will be more likely to follow the expectation strategy when making decisions for everyone and making decisions for him/herself

in a multiple-play condition, but will be more likely to follow heuristic/non-expectation strategies when making decisions for him/herself in a single-play condition. In recent years, eye-tracking technology has been successful in studying the complex cognitive activity involved in decision-making (Brandstätter & Körner, 2014; Liu et al., 2021; Liu et al., 2020; Sui et al., 2020; Wei & Li, 2015; Zhou et al., 2021). Therefore, in this study, we employed eye-tracking technology to compare each participant's process of information searching and processing during different tasks to test our hypotheses.

In the present study, we used three risky choice tasks, a decision-for-everyone (D-everyone) task, a decision-for-self-multiple-play (D-multiple) task, and a decision-for-self-single-play (D-single) task, to test our hypotheses. The purpose of the D-multiple task was that this task could be used as a control to determine whether the alternative-based strategy was adopted in the D-single task and the D-everyone task, given that researchers have demonstrated that individuals tend to follow the alternative-based strategy in the D-multiple task (Klos et al., 2005; Langer & Weber, 2001; Li, 2003; Sun et al., 2013).

The heuristic/non-expectation strategy and the expectation strategy make different predictions regarding participants' eye movement behaviors. In general, adopting these two different strategies will influence participants' scanpath patterns (Zhou et al., 2016). Specifically, adopting these strategies will also influence the participants' information search and processing in three ways: depth, complexity level, and direction (Su et al., 2013). Using the expectation strategy generally requires integrating all of the information available about the options, carrying out complex computations, and performing an alternative-based information search. In contrast, using a heuristic/non-expectation strategy requires selective use of information about the option, simple and ordinal comparisons, and a dimension-based information search pattern. Our working hypotheses for the present study were thus derived.

H_1: If decision-making in the D-everyone task is based on the same strategy as in the D-multiple task, but decision-making in the D-single task is based on a different strategy, then scanpath patterns in the D-everyone and D-multiple tasks will be similar, meanwhile contrasting with scanpath patterns in the D-single task.

H_2: If decision-making in the D-everyone task and in the D-multiple task is based on an expectation strategy, and decision-making in the D-single task is based on a heuristic/non-

expectation strategy, then the depth of information acquisition and the complexity of the information processing will be higher in the D-everyone and D-multiple tasks than in the D-single task, and the direction of the information search in the D-everyone and D-multiple tasks will be more alternative-based than that in the D-single task.

Moreover, if H_1 and H_2 are true, then we will also test the following hypothesis:

H_3: The eye-tracking measures (i.e., the depth of information acquisition, the complexity of the information processing, and the direction of the information search) mediate the effect of the task on participants' EV-consistent choice.

To further test whether individuals adopt the dimension-based or alternative-based strategies, researchers have commonly manipulated two types of factors by following the logic of double dissociation (Brandstätter et al., 2006; Liu et al., 2015; Rao et al., 2013; Zhang et al., 2018), which refers to a case where one factor affects performance on task A but not task B, while a second factor affects performance on task B but not task A. The rationale behind double dissociation is that if each task is process-pure, then a double dissociation provides evidence that the processes supporting tasks A and B are separate (Rao et al., 2013). The first type of factor is the overall difference between two options. For example, a decrease in the ratio between EVs was reported to be associated with a decrease in the proportion of correct predictions of expectation theories (Brandstätter et al., 2006; Brandstätter, Gigerenzer, & Hertwig, 2008). The second type of factor is the dimensional difference. For example, the difference in the outcome dimension was found to influence both the extreme choice proportion and the cognitive process in a dimension-based strategy (Brandstätter et al., 2006; Rao et al., 2013; Rao et al., 2011; Huang et al., 2021). Therefore, we also tested the following hypothesis, which concerns the impact of EV difference and the difference in the outcome dimension.

H_4: If decision-making in the D-everyone and D-multiple tasks is based on an expectation strategy and decision-making in the D-single task is based on a heuristic/non-expectation strategy, then the magnitude of the difference in the outcome dimension (hereafter referred to as outcome difference) should affect the participants' information processing (e.g., the depth of information acquisition) in the D-single task, while no effect of the EV difference should be observed; however, the magnitude of the EV difference should affect the participants' information processing (e.g., the complexity of the information processing) in the D-everyone and D-multiple tasks, while no effect

of the difference in the outcome dimension should be observed.

2 Methods

2.1 Participants

We used G*Power software (version 3.1.9.2) (Faul et al., 2007) to calculate the number of participants needed to achieve 95% power (1 – β) to detect the effect size of Cohen's f = 0.32 (based on Su et al. (2013)'s results). The necessary sample size was N = 46. A total of 52 college students (26 males, mean age = 22.8 ± 3.0 years) participated in the experiment. All had normal or corrected-to-normal vision and provided written informed consent prior to the experiment. Each participant was paid 120 yuan (RMB) in cash for participation, plus an additional amount (0–45 Yuan) that was determined by his or her performance during the experiment..Four participants were excluded from the analyses due to incomplete tracking data.

2.2 Apparatus

The participants' eye movements were monitored with an EyeLink 1000 tracker (SR Research, Canada) with the eye position sampled at 1,000 Hz. Participants viewed the stimuli with both eyes, but eye movement data was collected from only the right eye. The stimuli were presented on a 19-inch CRT monitor (with a refresh rate of 150 Hz) controlled by a Dell PC with a display resolution of 1,024 × 768 pixels. Although the eye-tracking system compensated for head movements, a chin rest located 60 cm away from the monitor was used to minimize head movements. Viewed from this distance, the screen subtended a visual angle of 37° horizontally and 28° vertically. Participants responded during the experiment by pressing a button on a Microsoft SideWinder gamepad.

2.3 Materials and Procedure

2.3.1 Stimuli

We constructed 108 pairs of two-payoff monetary risky options (see table S1 in supplemental materials: https://osf.io/yg67b/). All of the options involved gains only. The 108 pairs of options followed a 3 (EV difference: small, medium, large) × 3 (outcome difference: small, medium, large) design, and each condition contained 12 pairs of options. The EV difference referred to the expected value difference between the two options, and the 3 levels were 10, 30, and 60. Each

option had a best (maximum) outcome and a worst (minimum) outcome. The best outcomes of the two options had 3 levels—400, 300, and 200—while the worst outcomes were fixed at 100, resulting in 3 levels of outcome difference (i.e., 300, 200, and 100).

We adopted Brandstätter and Körner's (2014) task format to ensure that each piece of information was fixated on properly and that during fixation on a piece of information, peripheral identification of an adjacent interest area was impossible. Specifically, each value was surrounded by a quadratic frame whose edge was 3.50° long and 0.90° thick to provide a clear saccade target, minimizing peripheral identifiability of adjacent values (Bouma, 1970). The center-to-center distance between any two values was 5.95°, which ensured that the participants could not perceive or identify more than one value at a time without making an eye movement (Rayner, 2009). The outcomes were presented to the left of their respective probabilities (see Figure 1a). The position of the options was counterbalanced, i.e., the option with the higher EV was on either the top or the bottom.

2.3.2 Experimental Task

There were three risky choice tasks in the present study: a D-everyone task, a D-multiple task, and a D-single task. In the D-everyone task, participants were asked to choose the more optimal option out of the two options under the assumption that *their selection would be the final decision for everyone who was facing the same choice*—that is, everyone would be subject to the same choice but could receive different outcomes. In the D-multiple task, participants were asked to choose between the two options under the assumption that their selection would be applied a total of 100 times. In the D-single task, participants were asked to choose between the two options under the assumption that their selection would be applied only once to themselves.

To further incentivize their cooperation, participants were told that their payment would be determined by their performance during the experiment. In the D-single task and the D-multiple task, participants were told that at the end of the experiment one choice would be randomly selected to be played for real, with the relevant outcomes determined at random by a background program. For the D-single task, the selected choice would be played only once. For the D-multiple task, the selected choice would be played a total of 100 times—that is, one of their choices would be selected randomly and be played 100 times by the background program. The outcomes would be added to

their virtual winning accounts. The greater the total amount in their virtual accounts, the more they would be paid for the D-multiple task. In the D-everyone task, the participants were told that at the end of the experiment their payment would be based on whether they had selected the optimal options for everyone (note that participants were not asked to choose the option with the higher EV). Before the experiment began, the participants were informed that they would be paid after the entire three-task experiment was completed.

Each participant performed all three of the tasks, but they performed only one task on a given day, with an interval of no less than three days between any two subsequent tasks. The order of the tasks was counterbalanced across the participants. Visually identical experimental materials were used in all three of the tasks (see Figure 1 for examples).

2.3.3 Procedure

In each task, when the participants came into the laboratory, they were given instructions about the experiment and a brief description of the apparatus. The chair was then adjusted to make the participants feel comfortable. The eye tracker was calibrated at the beginning of the experiment and was calibrated again as needed (e.g., if the drift check failed). We used a 9-point calibration and validation procedure. The maximum error of validation was 0.5° in the visual angle. After the initial calibration, 4 practice trials were first presented to familiarize participants with the presentation pattern and task and to ensure that participants understood the experimental instructions before they began. The testing session contained a total of 108 trials, and the order of the trials was counterbalanced across participants. The 108 trials were divided into 3 blocks, each block containing 36 trials. Participants were permitted to take a 1–2 min break after finishing each block.

At the beginning of each trial, a fixation disc was presented at the center of the display; this also served as a drift check for the eye tracker. When fixation on that disc was registered, participants pressed a button to trigger the presentation of a choice stimulus. Then participants responded by pressing one of two other buttons. They were instructed to press the left button to indicate a decision for the top option or to press the right button to indicate a decision for the bottom option. There was no time limit, and the screen was cleared once the participant pressed a button. After the participant responded, a 1,000-ms interval (with a blank screen) followed before the next trial began (for details, see Figure 1). No feedback about the outcome of the selected option was provided until the entire

experiment was completed.

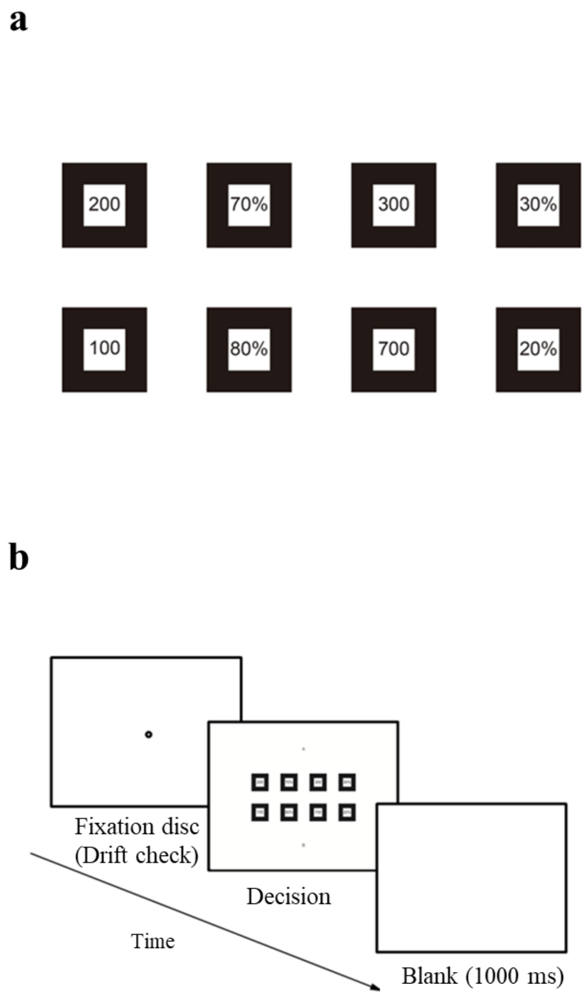

Figure 1. Diagram of stimulation paradigms. (a) Presentation of stimuli. (b) Trial procedure and timing.

2.4 Model prediction

We separately calculated the percentages of choices that were correctly predicted by (1) 3 expectation theories, including the EV theory (Bernoulli, 1738), the EU theory (von Neumann & Morgenstern, 1947), and the cumulative prospect theory (CPT) (Tversky & Kahneman, 1992), and (2) 3 heuristic/non-expectation theories , including the equate-to-differentiate (ETD) model (Li, 2004, 2016), the maximax heuristic (MH) (Brandstätter et al., 2006; Savage, 1951), and the tallying heuristic (TH) (Gigerenzer, 2004). We aimed to examine how well these theories

accounted for the participants' choices. The decision rules for these 6 models are summarized in Table 1.

Table 1 Decision rules of 6 models tested in predicting choice data

Model	Decision rule
Expected value (EV) theory	Calculate the sum of all weighted possible outcomes using the formula $\Sigma\, p_i \cdot x_i$. Choose the gamble with the highest weighted sum.
Expected utility (EU) theory	Calculate the sum of all weighted outcomes using the following formula: $\Sigma\, p_i \cdot u(x_i)$. Choose the gamble with the highest weighted sum. We assumed the utility function $u(x_i) = \log(x_i)$ in this study (Su et al., 2013).
Cumulative prospect theory (CPT)	Calculate the sum of all weighted outcomes using the following formula: $\Sigma\, \pi(p_i) \cdot v(x_i)$, $\pi(p_i) = p_i^{\gamma}/[\, p_i^{\gamma} + (1 - p_i)^{\gamma}]^{1/\gamma}$, $v(x_i) = x_i^{\alpha}$. Choose the gamble with the highest weighted sum. We estimated the values of parameters in individual level.
Equate-to-differentiate (ETD) model	Choose the gamble with more attractive gain on the dimension (best or worst payoff) with the greatest intradimensional utility difference. We assumed the utility function $u(x_i) = \log(x_i)$ in this study (Su et al., 2013).
Maximax heuristic (MH)	Choose the gamble with the highest monetary payoff.
Tallying heuristic (TH)	Give a tally mark to the gamble with (a) the higher minimum gain, (b) the higher maximum gain, (c) the lower probability of the minimum gain, and (d) the higher probability of the maximum gain. Select the gamble with the higher number of tally marks.

2.5 Data analysis for eye movement data

The eye movement data were analyzed with the Eyelink Data-Viewer software (SR Research, Canada). Fixations were defined as periods of relatively stable gaze between 2 saccades, but fixations shorter than 50 ms were excluded from analyses. Eight non-overlapping, identically sized (94×94 pixels) rectangular regions of interest (ROIs) were defined: 4 regions covered the payoffs of both options, and 4 regions covered the probabilities. We analyzed the scanpath patterns, the number of fixations on each ROI, the number of saccades between ROIs, and the mean fixation duration in all ROIs. To test whether participants showed different engagement in the 3 tasks, we also calculated the maximum pupil diameter in each trial. Because pupil size has been proved to be an index of task engagement (Hopstaken et al., 2015).

3 Results

Overall, 2 of the 15,552 trials were excluded from analyses because of eye-tracking failures.

A one-way repeated-measures ANOVA conducted on the maximum pupil diameter revealed no significant effect of task, $F(2, 94) = 2.20$, $p = 0.116$, $\eta^2_p = 0.05$, indicating that there was no difference in engagement among the 3 tasks.

3.1 Behavioral Results

3.1.1 Model prediction

A 3 (task) × 6 (model) repeated-measures ANOVA of the percentage of choices that were correctly predicted revealed a significant main effect of task, $F(2, 94) = 31.30$, $p < 0.001$, $\eta^2_p = 0.40$; a significant main effect of model, $F(5, 235) = 201.50$, $p < 0.001$, $\eta^2_p = 0.81$; and a significant interaction effect, $F(10, 470) = 31.20$, $p < 0.001$, $\eta^2_p = 0.40$. For the D-everyone task and the D-multiple task, pairwise comparisons showed that the percentages of choices that were correctly predicted by EV, EU and CPT were greater than the percentage predicted by other theories ($ts \geqslant 9.07$, $ps < 0.001$). For the D-single task, pairwise comparisons revealed that the percentages of choices that were correctly predicted by EV, EU, TH, EDT, and MH did not differ significantly from each other ($ts \leqslant 2.56$, $ps > 0.502$) (Figure 2).

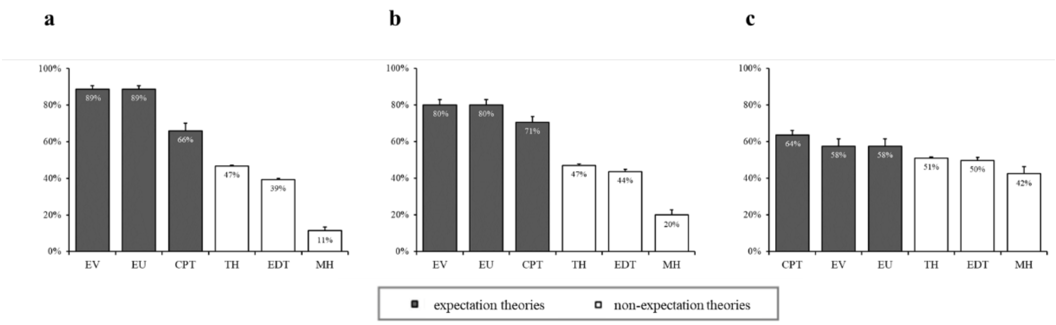

Figure 2. Percentage of choices that were correctly predicted by expectation theories and heuristic/non-expectation theories in (a) the D-everyone task, (b) the D-multiple task, and (3) the D-single task. Error bars represent standard errors of the means.

3.1.2 Decision time

Decision time was submitted to a natural log transformation to reduce skew. A one-way repeated-measures ANOVA conducted on the decision time revealed a significant main effect of

task, $F(2, 94) = 5.39$, $p = 0.006$, $\eta^2_p = 0.10$. Furthermore, a post hoc analysis revealed that the participants spent less time on the D-single task ($M = 8.75$, 95% CI = [8.63, 8.87]) than on the D-everyone task ($M = 8.91$, 95% CI = [8.79, 9.03], $t_{94} = 2.83$, $p = 0.015$) or on the D-multiple task ($M = 8.91$, 95% CI = [8.78, 9.04], $t_{94} = 2.86$, $p = 0.014$), with no significant difference between the latter two tasks ($t_{94} = 0.02$, $p = 1.000$).

The choice data revealed that neither the expectation theories nor the heuristic/non-expectation theories outperformed one another in predicting choices in the D-single task. The expectation theories, however, especially EV and EU theories, more accurately predicted choices than the heuristic/non-expectation theories did in the D-multiple and the D-everyone task (note that the heuristic/non-expectation theories did not claim that they could predict people's choice in the D-multiple and the D-everyone task). These results indicated that participants were more likely to follow the EV or EU theory in the D-everyone task and the D-multiple task than in the D-single task. Since more time is expected to be required to perform the expectation strategy than the heuristic/non-expectation strategy (Rao et al., 2011; Su et al., 2013), our decision time results also suggested that participants were more likely to make risky choices when following the expectation strategy in the D-everyone task and the D-multiple task than in the D-single task. The observed difference between the D-multiple task and the D-single task on the percentage of choices that were correctly predicted by theories, and the decision time replicated the main results obtained in Su et al's (2013) study.

3.2 Scanpath Analysis

To test whether scanpath patterns in the D-everyone task and the D-multiple task were different from patterns in the D-single task (H_1), we conducted a scanpath analysis using the

method described by Zhou et al. (2016). While viewing the screen during the cognition phase, individuals generate a relatively fixed "path". This path, which is characteristic of a given participant viewing a given pattern, is called a "scanpath" (Noton & Stark, 1971). Scanpath analysis is considered to be able to provide clear insights into the cognitive processes underlying behavioural decision-making theory (Ashby et al., 2016). An increasing similarity score indicates greater similarity between two scanpaths during risky choices. We used the ScanMatch toolbox (Cristino et al., 2010), which is based on the Needleman–Wunsch (N–W) algorithm (Needleman & Christian, 1970), to measure the similarities between the scanpaths. The parameters were set as Cristino et al.'s (2010) default values. Compared with traditional string-editing algorithms, the N–W algorithm can take the fixation length and fixation duration into account and can edit the setup of the stimulus regions of interest (ROIs) being compared. We calculated the intra-task and inter-task similarity scores separately for each participant.

The similarity score for the intra-task in the D-single task (M = 0.465, 95% CI = [0.454, 0.476]) was significantly lower than the score in the D-everyone task (M = 0.485, 95% CI = [0.468, 0.502], t_{47} = 2.66, p = 0.011, Cohen's d = 0.38) and in the D-multiple task (M = 0.484, 95% CI = [0.470, 0.498], t_{47} = 3.35, p = 0.002, Cohen's d = 0.48), with no significant difference between the latter two tasks, t_{47} = 0.09, p = 0.933, Cohen's d = 0.01 (Figure 3). This result indicated that the internal consistency of scanpath patterns in the D-single task was lower than that in the D-everyone and D-multiple tasks.

For all three of the tasks, similarity scores for the intra-task were higher than scores for the inter-task (ts ⩾ 5.56, ps < 0.001, Cohen's ds ⩾ 0.80), indicating that scanpath patterns of the three tasks differed from each other. Further analysis revealed that the similarity score for the

inter-task between the D-everyone and D-multiple tasks ($M = 0.454$, 95% CI = [0.439, 0.469]) was higher than that between the D-everyone and D-single tasks ($M = 0.432$, 95% CI = [0.417, 0.446], $t_{47} = 3.46$, $p = 0.001$, Cohen's $d = 0.50$) and that between the D-multiple and D-single tasks ($M = 0.434$, 95% CI = [0.420, 0.449], $t_{47} = 3.53$, $p < 0.001$, Cohen's $d = 0.51$). However, the difference between the latter two was not significant ($t_{47} = 0.46$, $p = 0.645$, Cohen's $d = 0.07$). These results suggested that scanpath patterns in the D-everyone task and D-multiple task were more similar than those in the D-single task.

Taken together, the similarity score results indicated that although scanpath patterns in the three tasks differed from each other, scanpath patterns in the D-everyone task and D-multiple task were more similar than those in the D-single task, thus supporting H_1.

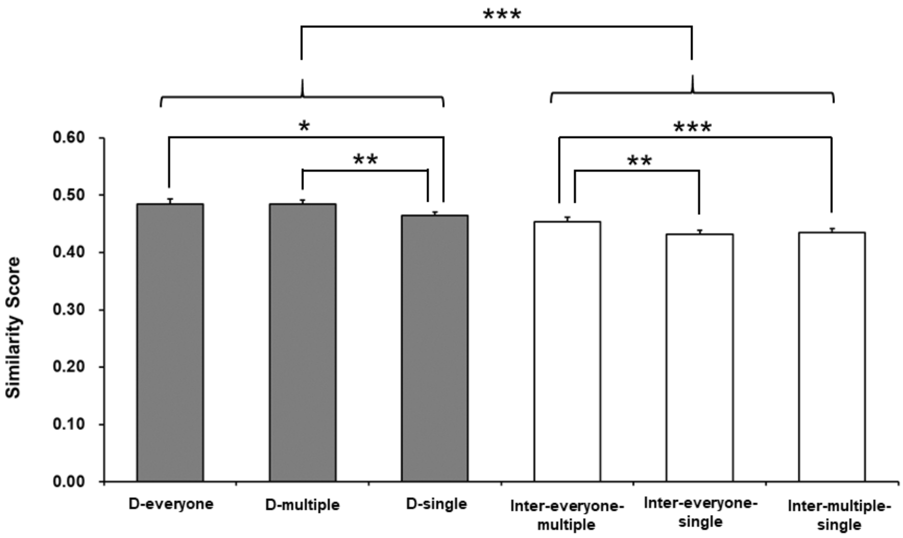

Figure 3. Similarity scores for the intra-task and the inter-task in the D-everyone, D-multiple, and D-single tasks. Error bars represent standard errors of the means. $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$.

3.3 Eye-Tracking Measures

To test H_2, we used the following eye-tracking measures. (1) Consistent with previous studies (Payne & Braunstein, 1978; Su et al., 2013), we used the percentage of total information searched

(PTIS) in each task to measure the depth of information acquisition. (2) We used the mean fixation duration (MFD) to measure the levels of complexity of information processing in decision making research. The duration of a single fixation increases with an increase in the level of complexity of information processing (Velichkovsky et al., 2002). (3) We used the alternative-based versus dimension-based SM index (SMI) and the proportion of the saccades between the two best outcomes (PSTB) to quantify the degree to which the direction of a search was alternative-based or dimension-based (Böckenholt & Hynan, 1994; Pachur et al., 2013; Su et al., 2013). The predominance of alternative-based saccades increases with an increasing SMI value (Liu et al., 2021; Su et al., 2013). If the participants make risky choices following an expectation strategy, the PTIS, MFD, and SMI should be higher, whereas the PSTB should be lower, compared with following a heuristic/non-expectation strategy.

We conducted 3 (task) × 3 (EV difference) × 3 (outcome difference) repeated-measure ANOVA with the PTIS, MFD, SMI, and PSTB as dependent variables separately. The results are presented in Table 2.

The results revealed a significant main effect of task for all four eye-tracking measures (Table 2). Post-hoc analyses revealed that (1) the PTIS was significantly lower in the D-single task (M = 92.8%, 95% CI = [91.0%, 94.5%]) than in the D-multiple task (M = 95.6%, 95% CI = [93.8%, 97.3%], t_{94} = 2.88, p = 0.014) and in the D-everyone task (M = 95.3%, 95% CI = [93.7%, 97.1%], t_{94} = 2.59, p = 0.029) (Figure 4a); (2) the MFD was shorter in the D-single task (M = 224 ms, 95% CI = [211, 238] ms) than in the D-multiple task (M = 246 ms, 95% CI = [233, 259] ms, t_{94} = 3.48, p = 0.002) and in the D-everyone task (M = 252 ms, 95% CI = [239, 265] ms, t_{94} = 4.45, p < 0.001) (Figure 4b); (3) the SMI was significantly lower in the D-single task (M = 0.79, 95% CI =

[0.55, 1.03]) than in the D-multiple task ($M = 1.17$, 95% CI = [0.93, 1.40], $t_{94} = 3.14$, $p = 0.006$) and in the D-everyone task ($M = 1.26$, 95% CI = [1.02, 1.50], $t_{94} = 3.92$, $p < 0.001$) (Figure 4c); (4) the PSTB was significantly higher in the D-single task ($M = 2.0\%$, 95% CI = [1.7%, 2.2%]) than in the D-multiple task ($M = 1.5\%$, 95% CI = [1.3%, 1.8%], $t_{94} = 3.24$, $p = 0.005$) and in the D-everyone task ($M = 1.4\%$, 95% CI = [1.2%, 1.7%], $t_{94} = 4.00$, $p < 0.001$) (Figure 4d). No significant differences between the D-multiple task and D-everyone task were observed for the four eye-tracking measures.

Table 2 Summary of repeated measures ANOVA on PTIS, MFD, SMI, and PSTB across task, EV difference, and outcome difference

Variable	*df*	*F*	*p*	η^2_p
PTIS				
Task	2, 94	5.02	0.008	0.10
ED	2, 94	0.19	0.827	0.00
OD	2, 94	3.68	0.029	0.07
Task × ED	4, 188	1.03	0.396	0.02
Task × OD	4, 188	2.96	0.021	0.06
ED × OD	4, 188	0.57	0.686	0.01
Task × ED × OD	8, 376	0.93	0.488	0.02
MFD				
Task	2, 94	10.94	< 0.001	0.19
ED	2, 94	22.75	< 0.001	0.33
OD	2, 94	2.93	0.058	0.06
Task × ED	4, 188	5.78	< 0.001	0.11
Task × OD	4, 188	1.22	0.304	0.03
ED × OD	4, 188	3.65	0.007	0.07
Task × ED × OD	8, 376	1.79	0.079	0.04
SMI				
Task	2, 94	8.59	< 0.001	0.16
ED	2, 94	0.72	0.487	0.02
OD	2, 94	0.32	0.725	0.01
Task × ED	4, 188	1.60	0.177	0.03
Task × OD	4, 188	0.16	0.961	0.00
ED × OD	4, 188	6.84	< 0.001	0.13
Task × ED × OD	8, 376	1.20	0.298	0.03
PSTB				
Task	2, 94	9.03	< 0.001	0.16
ED	2, 94	1.49	0.230	0.03
OD	2, 94	5.14	0.008	0.10

Task × ED	4, 188	0.96	0.433	0.02
Task × OD	4, 188	0.86	0.491	0.02
ED × OD	4, 188	0.33	0.860	0.01
Task × ED × OD	8, 376	1.30	0.244	0.03

Note. ED: EV difference; OD: outcome difference.

The ANOVA revealed a significant interaction effect between task and outcome difference for the PTIS (Table 1). Simple effect analysis showed that in the D-single task, the PTIS for the large outcome difference level was significantly lower than that for the small outcome difference level ($t_{277} = 4.04$, $p = 0.002$). No significant difference in the PTIS was found for the D-everyone task or the D-multiple task. The finding that the outcome difference affected the PTIS only in the D-single task supported H_4.

The ANOVA also revealed a significant interaction effect between task and EV difference for the MFD (Table 1). Simple effect analysis with Bonferroni adjustments showed that in the D-everyone task, the MFD for the large EV difference level was significantly shorter than that for medium ($t_{281.9} = 7.65$, $p < 0.001$) and small ($t_{281.9} = 3.28$, $p = 0.031$) EV difference levels[1]. No significant difference in the MFD was found for the D-multiple task or the D-single task. The finding that the EV difference affected the MFD only in the D-everyone task also supported H_4.

[1] Unexpectedly, the MFD in the medium EV difference level was significantly longer than that in the small EV difference level. This unexpected result might have been caused by the computational difficulty in the small EV difference level being lower than that in the medium EV difference level. We speculated that the computational difficulty for calculating the expected value of options with the probability of 10% (or options with an outcome of 100 RMB) was lower than that of options with other probabilities (or options with other outcomes). A post hoc observation found that the options in the small EV difference level consisted of more options with low computational difficulty than those in the medium EV difference level (26 options with a low computational difficulty in the small EV difference level, 18 options in the medium EV difference level). Our speculation was partially supported by the finding that there was no significant difference in MFD between the small EV difference level and the medium EV difference level in the D-everyone task ($F(1, 47) = 1.79$, $p = 0.187$, $\eta^2_p = 0.04$) after controlling for computational difficulty.

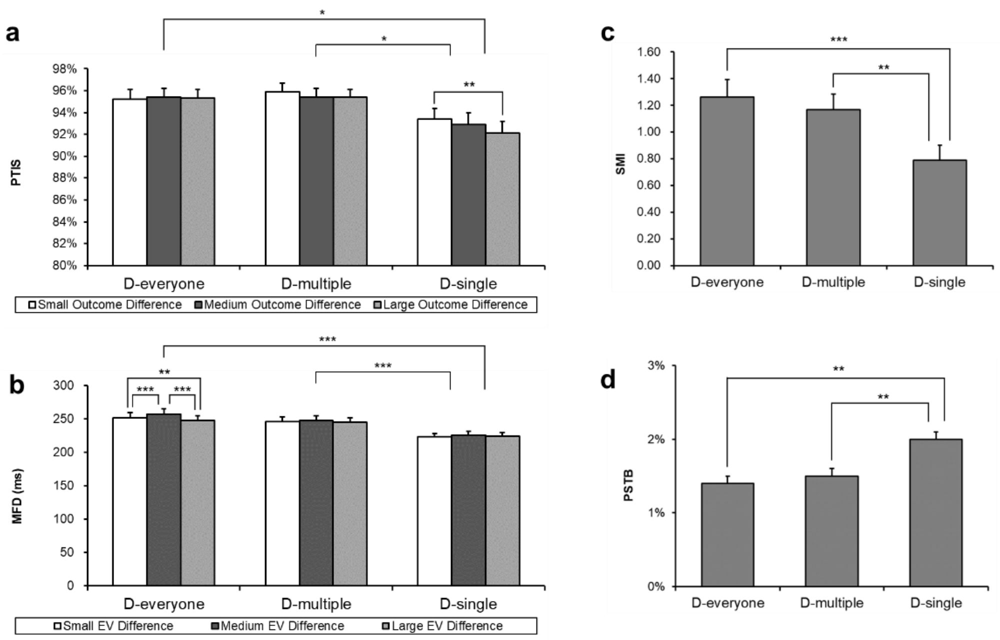

Figure 4. Results of eye-tracking measures. (a) The percentage of total information searched (PTIS). (b) The mean fixation duration (MFD). (c) The alternative-based versus dimension-based search measure index (SMI). (d) The proportion of the saccades between the two best outcomes (PSTB). Error bars represent standard errors of the means. $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$.

The above results indicated that participants performed a lower depth of information acquisition, a lower complexity level of information processing, and a more dimension-based information search in the D-single task, compared with the D-multiple task and D-everyone task. These results suggested that people are more likely to use a heuristic/non-expectation strategy in the D-single task when making a choice between risky options. Making risky choices in the D-everyone task and D-multiple task was more consistent with the weighting and adding process assumed by an expectation strategy.

3.4 Mediation Analysis

To test H_3, we examined whether the eye-tracking measures mediated the effect of task on choice. We used the EV-consistent choice (i.e., the choice that was correctly predicted by the EV strategy) as the dependent variable. We used the parallel multiple mediator model in MEMORE

(Montoya & Hayes, 2017) for SPSS to test the effect of task on the EV-consistent choice through several parallel mediators. The parallel multiple mediator model allows correlation between mediators and can estimate the indirect effect of each mediator after controlling for the effect of other proposed mediators in the model. Since only the two-condition within-participant mediation can be estimated using MEMORE, we conducted two mediation analyses. That is, we examined the effect of two tasks (either D-single task vs. D-everyone task or D-single task vs. D-multiple task) on the EV-consistent choice through all aforementioned eye-tracking measures. Ninety-five percent confidence intervals were generated based on 5,000 bootstrap samples and are reported in brackets for each result below.

D-single task vs. D-everyone task. Mediation analysis results are shown in Figure 5a. Mediation analysis revealed significant indirect effects of task on the EV-consistent choice through the four eye-tracking measures: the PTIS (a_1b_1 = 0.008 [0.005, 0.012]), MFD (a_2b_2 = 0.021 [0.013, 0.028]), SMI (a_3b_3 = 0.014 [0.009, 0.018]), and PSTB (a_4b_4 = 0.006 [0.004, 0.008]). The total effect of task on the EV-consistent choice was significant (c = 0.31 [0.29, 0.32], p < 0.001), and the direct effect of task on the EV-consistent choice (controlling for the influence of the mediators) remained significant (c' = 0.26 [0.24, 0.28], p < 0.001), indicating that task still accounted for variance in the EV-consistent choice over and above the effects of mediators (Figure 5a). Pairwise contrasts of each indirect effect showed that the MFD made a stronger contribution to the EV-consistent choice than the PTIS ($a_2b_2 - a_1b_1$ = 0.013 [0.004, 0.020]) and PSTB ($a_2b_2 - a_4b_4$ = 0.015 [0.007, 0.023]), and that the SMI made a stronger contribution to the EV-consistent choice than the PSTB ($a_3b_3 - a_4b_4$ = 0.008 [0.002, 0.013]). However, the MFD and SMI did not differ from each other ($a_3b_3 - a_2b_2$ = −0.007 [−0.016, 0.002]).

D-single task vs. D-multiple task. Mediation analysis results are shown in Figure 5b. Mediation analysis revealed significant indirect effects of task on the EV-consistent choice through the four eye-tracking measures: the PTIS (a_1b_1 = 0.010 [0.006, 0.014]), MFD (a_2b_2 = 0.030 [0.023, 0.036]), SMI (a_3b_3 = 0.012 [0.008, 0.016]), and PSTB (a_4b_4 = 0.003 [0.001, 0.005]). The total effect of task on the percentage of the EV-consistent choice was significant (c = 0.22 [0.21, 0.24], $p < 0.001$), and the direct effect of task on the EV-consistent choice (controlling for the influence of the mediators) remained significant (c' = 0.17 [0.15, 0.19], $p < 0.001$), indicating that task still accounted for variance in the EV-consistent choice over and above the effects of mediators (Figure 5b). Pairwise contrasts of each indirect effect showed that the MFD made a stronger contribution to the EV-consistent choice than the other three measures ($a_2b_2 - a_1b_1$ = 0.020 [0.012, 0.028], $a_2b_2 - a_3b_3$ = 0.018 [0.010, 0.026], $a_2b_2 - a_4b_4$ = 0.027 [0.020, 0.034]). Pairwise contrasts of each indirect effect also showed that the PTIS and SM index made a stronger contribution to the percentage of the EV-consistent choice than PSTB ($a_1b_1 - a_4b_4$ = 0.007 [0.002, 0.011], $a_3b_3 - a_4b_4$ = 0.009 [0.004, 0.014]).

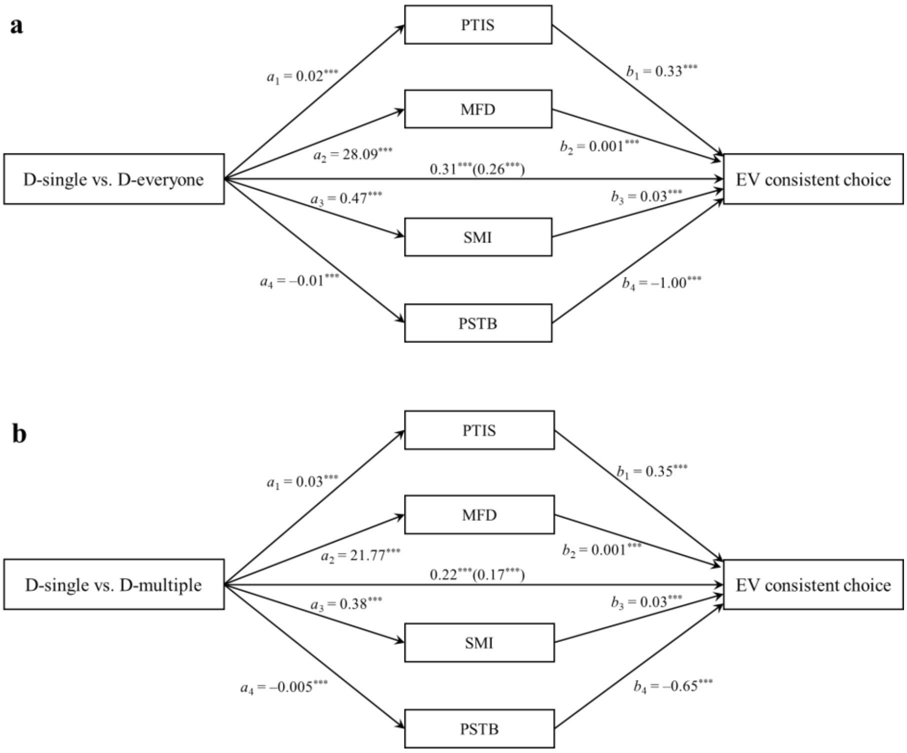

Figure 5. Results of mediation analysis. (a) Mediating effect of eye-tracking measures on the relationship between task (D-single task vs. D-everyone task) and the percentage of the EV-consistent choice. (b) Mediating effect of eye-tracking measures on the relationship between task (D-single task vs. D-multiple task) and EV-consistent choice. $^{***}p < 0.001$.

In accordance with our hypothesis, mediation analysis results found that the eye-tracking measures mediated the relation between task and the EV-consistent choice. Among all eye-tracking measures, the MFD (an index of the complexity level of the information processing) and SMI (an index of a search is alternative-based or dimension-based) appeared to be the most important mediators of the relationship between task and EV-consistent choice. Compared with the D-single task, in the D-everyone task and D-multiple task, the depth of information acquisition and the complexity level of information processing were higher, and the direction of information search was more alternative-based, thus leading to an increase in the EV-consistent choice.

4 Discussion

The debate about whether making a risky choice is based on the expectation strategy or on a

heuristic/non-expectation strategy is currently unresolved (Pachur et al., 2013). From the perspective of process testing, the present study used an eye-tracking technique to explore the boundaries of the expectation theories and heuristic/non-expectation models.

An overview of our behavioral and eye movement analysis results is provided in Figure 6. Behavioral results revealed that, compared with the D-single task, participants selected more choices correctly predicted by EV and EU theories, and took a longer time to make a decision in the D-everyone and D-multiple tasks. Furthermore, eye movement measurements revealed that: (1) the scanpath patterns of the D-everyone task and D-multiple task were similar, but different from those of the D-single task; (2) the depth of information acquisition in the D-everyone task and D-multiple task was higher than that in the D-single task, and the outcome difference affected the depth of information acquisition only in the D-single task; (3) the level of complexity of information processing in the D-everyone task and D-multiple task was higher than that in the D-single task, while the magnitude of the EV difference affected the complexity of information processing only in the D-everyone task; (4) the direction of information search in the D-everyone task and D-multiple task was more alternative-based than that in the D-single task; and (5) the eye-tracking measures mediated the relationship between task and the EV-consistent choice. In summary, both behavioral and eye movement results supported our hypotheses that participants were likely to follow an expectation strategy in the D-everyone and D-multiple tasks, while being likely to follow a heuristic/non-expectation strategy in the D-single task.

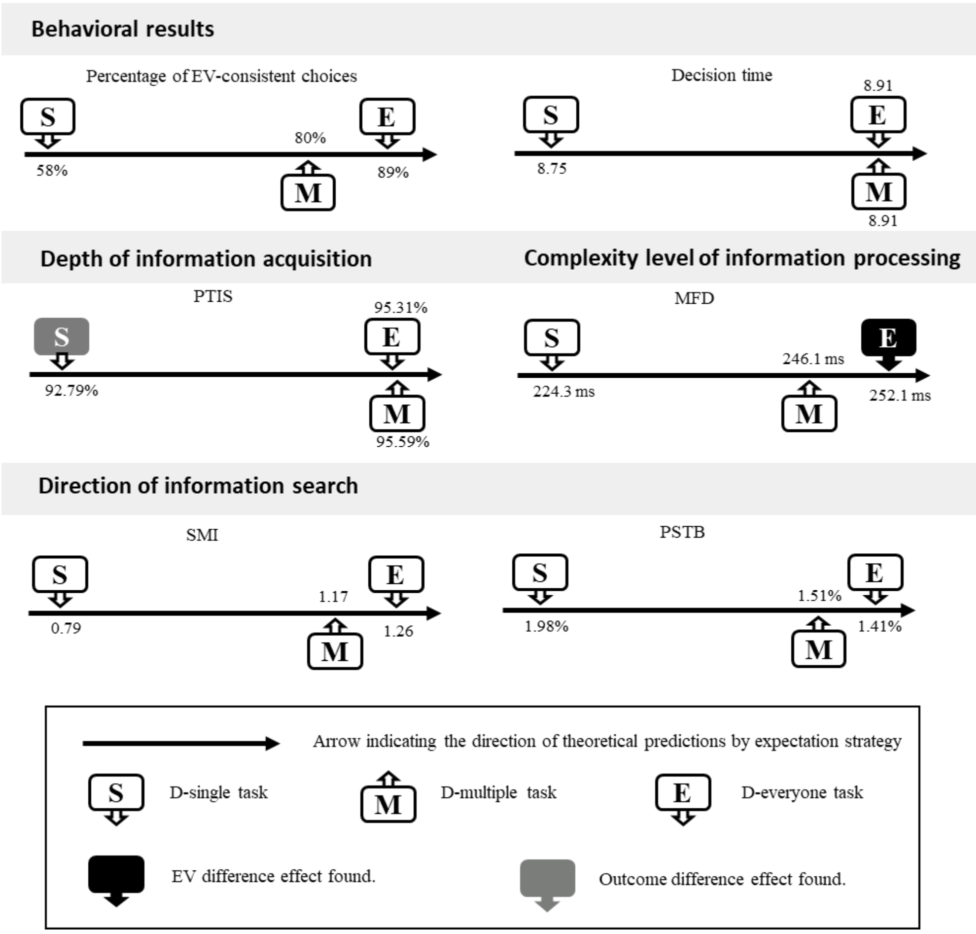

Figure 6. Overview of behavioral and eye movement results.

Returning to the question raised in the Introduction section about why mainstream theorists have not abandoned the expectation rule, we think that this may possibly be because these theorists were reinforced to stick to these expectation theories by two beliefs: One belief is that expectation theories should work well for everyone, while the other belief is that these theorists by default believe that every one is a subset of everyone. These two beliefs led to the view that expectation theories could eventually capture every one's decision-making under risk.

Regarding the first belief, the present study provided the first eye-tracking evidence from a perspective of process testing to support the idea that the elegant formulas of expectation theories work when every individual makes risky choices for everyone. This concept was supported by our

findings in the D-multiple task, given that multiple playing times and multiple individuals could be seen as similar forms under the guidance of the law of large numbers. By focusing on either the decision process or the decision outcome, the primary goal of previous research was to investigate whether expectation theories could describe an individual's risky choice when making decisions for self (Brandstätter & Körner, 2014; Li, 2004; Pachur et al., 2013; Rao et al., 2013; Sun et al., 2013) or making decisions for other people (Beisswanger et al., 2003; Hsee & Weber, 1997; Mengarelli et al., 2014; Polman, 2012). However, to our knowledge, few studies have attempted to explore whether expectation theories work for risky decision-making for everyone. Our results showed that our participants, as if following the advice of the dean in the apocryphal story about Raiffa, appeared to perform a weighting and adding process to reach decisions when making a decision for everyone.

Regarding the second belief, mainstream theorists take it for granted that expectation theories, which work ideally for decisions for everyone (full set), must also work for decisions for an individual (subset). However, the behavioral and eye movement evidence from the current study on the self–everyone discrepancy revealed that the decision for self was different from the decision for everyone. A decision for everyone is one thing, but a decision for every one is another. Therefore, a theory that works well for decision-making for everyone will not necessarily work for decision-making for every one. The evidence for this discrepancy between everyone and every one, which was first reported in our study, implied that the reason why expectation theories do not work could be that a default compatibility between the full set (everyone) and the subset (every one) does not exist.

When reviewing the development of the EV strategy proposed by Pascal and Fermat, we

noticed two points worth noting. First, the original intent of Pascal and Fermat was to provide an optimal solution to the problem of points that had been raised by a gambler, making the solution applicable to everyone (not Pascal and Fermat themselves) faced with the same problem (Durrett, 2010; Vinod & Reagle, 2005). In keeping with Pascal and Fermat's intention, our behavioural and eye movement results in the D-everyone and D-multiple tasks fit the predictions of the EV strategy better than the results in the D-single task, suggesting that Pascal and Fermat's EV theory could indeed capture the nature of risky choices when individuals make decisions for everyone. Second, EV strategy correctly predicts the observed behavior and eye movement in the D-everyone task (the predictive accuracy of the EV theory in the D-everyone task reached 89%) and appears to be predictively inadequate in the D-single task (the predictive accuracy of the EV theory in the D-single task reached 58%), implying that the simpler EV strategy is adequate for describing an individual's decision-making for everyone.

The findings of the present study may shed light on the general issue of classifying risky decision-making theories. Since the foundation of the EV theory, the family of expectation theories has been modified repeatedly due to their poor ability to cope with empirical challenges, such as the St. Petersburg paradox, Allais paradox, or other newer paradoxes (Birnbaum, 2008). Traditional EV theory was transformed into expected utility (EU) theory in its earliest Bernoullian form, to the axiomatic form of von Neumann and Morgenstern's EU theory (von Neumann & Morgenstern, 1947), to Savage's subjective EU theory (Savage, 1954), to the weighted utility model (Edwards, 1962), to the rank-dependent utility model (Quiggin, 1982), to the rank- and sign-dependent utility model (Luce & Fishburn, 1991), to the third-generation prospect theory (Schmidt et al., 2008), and so on. The soul of these modifications is reflected in Tversky's

comment, “the issue, therefore, is not whether choice can be described as a maximization, but rather which function is being maximized” (see Li, 2016, pp. 68–69). However, such non-stop modification seems to imply that no existing expectation theory has satisfactorily described an individual’s actual risky choice. Possibly as a result of the persistently poor performance of expectation theories in describing an individual’s actual risky choice, theorists in decision-making have had no other choice but to classify the existing decision-making theories into two categories: normative theories, which can be seen as a standard that defines thinking that is best for achieving the thinker’s goals, and descriptive theories, which are theories about how people normally think (Baron, 2008). In other words, theorists have had to tolerate the existence of expectation theories as normative theories without requesting them to describe an individual’s actual risky choice as descriptive theories. Unlike the commonly accepted concept of normative theories, our findings suggested that EV theory, which is categorized as a normative theory, can describe an individual’s actual risky choice when individuals make decisions for everyone.

Perhaps, if a theory is needed to describe an individual’s actual risky choice for self, the urgent need is not to modify expectation theories but to develop and improve heuristic/non-expectation decision theories from a non-compensatory and non-expectation-maximization perspective.

5 Conclusion

In the present study, we used an eye-tracking technique to examine participants’ behavior and eye-tracking measures in the D-everyone task, the D-multiple task and the D-single task. WE found that (1) compared with the D-single task, participants selected more choices correctly

predicted by EV and EU theories, and took a longer time to make a decision in the D-everyone and D-multiple tasks; (2) the scanpath patterns of the D-everyone task and D-multiple task were similar but different from those of the D-single task; (3) The eye-tracking measures mediated the relationship between the task and the EV-consistent choice; (4) the outcome difference affected individual's information process in the D-single task and the EV difference affected individual's information process in the D-everyone task. Our findings suggest that expectation-maximization-based theories could capture the choice of an individual when making decisions for everyone and for self in a multiple-play condition but could not capture the choice of an individual when making decisions for self in a single-play condition. These findings contribute to an improved understanding of the boundaries of expectation-maximization-based theories and those of heuristic/non-expectation models and may also shed light on the general issue of the classification of risky decision-making theories.

References

Ashby, N. J. S., Johnson, J. G., Krajbich, I., & Wedel, M. (2016). Applications and innovations of eye-movement research in judgment and decision making. *Journal of Behavioral Decision Making, 29*, 96–102.

Böckenholt, U., & Hynan, L. S. (1994). Caveats on a process-tracing measure and a remedy. *Journal of Behavioral Decision Making, 7*(2), 103–117.

Baron, J. (1986). Tradeoffs among reasons for action. *Journal for the Theory of Social Behavior, 16*(2), 173–195.

Baron, J. (2008). *Thinking and deciding* (4th ed.). Cambridge, England: Cambridge University Press.

Bazerman, M. H., Tenbrunsel, A. E., & Wade-Benzoni, K. (1998). Negotiating with yourself and losing: Making decisions with competing internal preferences. *Academy of Management Review, 23*(2), 225–241.

Beisswanger, A. H., Stone, E. R., Hupp, J. M., & Allgaier, L. (2003). Risk taking in relationships: Differences in deciding for oneself versus for a friend. *Basic and Applied Social Psychology, 25*(2), 121–135.

Bernoulli, D. (1738). Specimen theoriae novae de mensura sortis [Exposition of a new theory on the measurement of risk]. *Commentarii Academiae Scientiarum Imperialis Petropolitanae, 5*, 175–192.

Birnbaum, M. H. (2008). New tests of cumulative prospect theory and the priority heuristic: Probability-outcome tradeoff with branch splitting. *Judgment and Decision Making, 3*(4), 304–316.

Bouma, H. (1970). Interaction effects in parafoveal letter recognition. *Nature, 226*, 177–178.

Brandstätter, E., Gigerenzer, G., & Hertwig, R. (2006). The priority heuristic: Making choices without trade-offs. *Psychological Review, 113*(2), 409–432.

Brandstätter, E., Gigerenzer, G., & Hertwig, R. (2008). Risky choice with heuristics: Reply to Birnbaum (2008), Johnson, Schulte-Mecklenbeck, & Willemsen (2008) and Rieger & Wang (2008). *Psychological Review, 115*(1), 281–289.

Brandstätter, E., & Körner, C. (2014). Attention in risky choice. *Acta Psychologica, 152,* 166–176.

Cristino, F., Mathot, S., Theeuwes, J., & Gilchrist, I. D. (2010). ScanMatch: A novel method for comparing fixation sequences. *Behavior Research Methods, 42*(3), 692–700.

DeKay, M. L., Hershey, J. C., Spranca, M. D., Ubel, P. A., & Asch, D. A. (2006). Are medical treatments for individuals and groups like single-play and multiple-play gambles? *Judgment and Decision Making, 1*(2), 134–145.

DeKay, M. L., & Kim, T. G. (2005). When things don't add up the role of perceived fungibility in repeated-play decisions. *Psychological Science, 16*(9), 667–672.

Durrett, R. (2010). *Probability: Theory and Examples* (4th ed.). New York: Cambridge University Press.

Edwards, W. (1962). Subjective probabilities inferred from decisions. *Psychological Review, 69*(2), 109–135.

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175–191.

Gigerenzer, G. (2004). Fast and frugal heuristics: The tools of bounded rationality. In D. Koehler & N. Harvey (Eds.), *Blackwell handbook of judgment and decision making* (pp. 62–88). Oxford, United Kingdom: Blackwell.

Hopstaken, J. F., van der Linden, D., Bakker, A. B., & Kompier, M. A. (2015). A multifaceted investigation of the link between mental fatigue and task disengagement. *Psychophysiology, 52*(3), 305–315.

Hsee, C. K., & Weber, E. U. (1997). A fundamental prediction error: Self–others discrepancies in risk preference. *Journal of Experimental Psychology: General, 126*(1), 45–53.

Huang, Y., Shen, S., Yang, S., Kuang, Y., Li, Y., & Li, S. (2021). Asymmetrical property of the subproportionality of weighting function in prospect theory: Is it real and how can it be achieved? *Symmetry, 13*(10), 1928.

Kahneman, D., Sibony, O., & Sunstein, C. R. (2021). *Noise: A Flaw in Human Judgment*. Glasgow, Scotland:

William Collins Publishers.
Kahneman, D., & Lovallo, D. (1993). Timid choices and bold forecasts: A cognitive perspective on risk taking. *Management Science, 39*(1), 17–31.
Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the Econometric Society, 47*(2), 263–292.
Klos, A., Weber, E. U., & Weber, M. (2005). Investment decisions and time horizon: Risk perception and risk behavior in repeated gambles. *Management Science, 51*(12), 1777–1790.
Langer, T., & Weber, M. (2001). Prospect theory, mental accounting, and differences in aggregated and segregated evaluation of lottery portfolios. *Management Science, 47*(5), 716–733.
Li, S. (2003). The role of expected value illustrated in decision-making under risk: Single-play vs multiple-play. *Journal of Risk Research, 6*(2), 113–124.
Li, S. (2004). A behavioral choice model when computational ability matters. *Applied Intelligence, 20*, 147–163.
Li, S. (2016). *An equate-to-differentiate way of decision-making*. Shanghai, China: East China Normal University Press.
Liu, H. Z., Jiang, C. M., Rao, L. L., & Li, S. (2015). Discounting or priority: Which rule dominates the intertemporal choice process? *Acta Psychologica Sinica, 47*(4), 522–532.
Liu, H. Z., Wei, Z. H., & Li, P. (2021). Influence of the manner of information presentation on risky choice. *Frontiers in Psychology, 12,* 605206.
Liu, H. Z., Zhou, Y. B., Wei, Z. H., & Jiang, C. M. (2020). The power of last fixation: Biasing simple choices by gaze-contingent manipulation. *Acta Psychologica, 208*, 103106.
Luce, R. D., & Fishburn, P. C. (1991). Rank-and sign-dependent linear utility models for finite first-order gambles. *Journal of Risk and Uncertainty*, *4*(1), 29–59.
Mengarelli, F., Moretti, L., Faralla, V., Vindras, P., & Sirigu, A. (2014). Economic decisions for others: An exception to loss aversion law. *PLoS One, 9*(1), e85042.
Montoya, A. K., & Hayes, A. F. (2017). Two-condition within-participant statistical mediation analysis: A path-analytic framework. *Psychological Methods, 22*(1), 6–27.
Needleman, S. B., & Christian, C., D. (1970). A general method applicable to the search for similarities in the amino acid sequence of two proteins. *Journal of Molecular Biology*, *48*(3), 443–453.
Noton, D., & Stark, L. (1971). Scanpaths in eye movements during pattern perception. *Science, 171*, 308–311.
Pachur, T., Hertwig, R., Gigerenzer, G., & Brandstätter, E. (2013). Testing process predictions of models of risky choice: A quantitative model comparison approach. *Frontiers in Psychology, 4*, 646.
Payne, J. W., & Braunstein, M. L. (1978). Risky choice: An examination of information acquisition behavior. *Memory & Cognition, 6*(5), 554–561.
Polman, E. (2012). Self–other decision making and loss aversion. *Organizational Behavior and Human Decision Processes, 119*(2), 141–150.
Popovic, N. F., Pachur, T., & Gaissmaier, W. (2019). The gap between medical and monetary choices under risk persists in decisions for others. *Journal of Behavioral Decision Making, 32*(4), 388–402.
Quiggin, J. (1982). A theory of anticipated utility. *Journal of Economic Behavior & Organization, 3*(4), 323–343.
Rao, L. L., Liu, X. N., Li, Q., Zhou, Y., Liang, Z. Y., Sun, H. Y., . . . Li, S. (2013). Toward a mental arithmetic process in risky choices. *Brain and Cognition, 83*(3), 307–314.
Rao, L. L., Zhou, Y., Xu, L., Liang, Z. Y., Jiang, T., & Li, S. (2011). Are risky choices actually guided by a compensatory process? New insights from fMRI. *PLoS One, 6*(3), e14756.
Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *The Quarterly Journal of Experimental Psychology*, *62*(8), 1457–1506.

Savage, L. J. (1951). The theory of statistical decision. *Journal of the American Statistical Association, 46*(253), 55–67.

Savage, L. J. (1954). *The foundations of statistics*. New York: Wiley.

Schmidt, U., Starmer, C., & Sugden, R. (2008). Third-generation prospect theory. *Journal of Risk and Uncertainty, 36*(3), 203–223.

Su, Y., Rao, L. L., Sun, H. Y., Du, X. L., Li, X., & Li, S. (2013). Is making a risky choice based on a weighting and adding process? An eye-tracking investigation. *Journal of Experimental Psychology: Learning, Memory and Cognition, 39*(6), 1765–1780.

Sui, X. Y., Liu, H. Z., & Rao, L. L. (2020). The timing of gaze-contingent decision prompts influences risky choice. *Cognition, 195*, 104077.

Sun, H. Y., Rao, L. L., Zhou, K., & Li, S. (2013). Formulating an emergency plan based on expectation-maximization is one thing, but applying it to a single case is another. *Journal of Risk Research, 17*(7), 785–814.

Sun, H. Y., Su, Y., Zhou, K., & Li, S. (2011). From multiple-play to single-play in risky decision making: Quantitative change or qualitative change?. *Advances in Psychological Science, 19*(10), 1417–1425.

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty, 5*(4), 297–323.

van der Stoep, S. W., & Seifert, C. M. (1994). Problem solving, transfer, and thinking. In P. R. Pintrich, D. R. Brown, & C. E. Weinstein (Eds.), *Student motivation, cognition, and learning: Essays in honor of Wilbert J. McKeachie* (pp. 27–49). Hillstale, NJ: Erlbaum.

Velichkovsky, B. M., Rothert, A., Kopf, M., Dornhöfer, S. M., & Joos, M. (2002). Towards an express-diagnostics for level of processing and hazard perception. *Transportation Research Part F: Traffic Psychology and Behaviour, 5*(2), 145–156.

Vinod, H. R. D., & Reagle, D. (2005). *Preparing for the worst: Incorporating downside risk in stock market investments*. Hoboken, NJ: Wiley.

von Neumann, J., & Morgenstern, O. (1947). *Theory of games and economic behavior*. Princeton, NJ: Princeton University Press.

Wei, Z. H., & Li, X. Decision process tracing: Evidence from eye-movement data. *Advances in Psychological Science, 23*(12), 2029–2041.

Zhang, Y. Y., Zhou, L., You, X., Li, S., & Liang, Z. Y. (2018). Is intertemporal decision-making based on a temporal discounting process? Evidence from double-dissociation paradigm. *Scientia sinica Vitae, 48*(5), 592–608.

Zhou, L., Zhang, Y. Y., Wang, Z. J., Rao, L. L., Wang, W., Li, S., . . . Liang, Z. Y. (2016). A scanpath analysis of the risky decision-making process. *Journal of Behavioral Decision Making, 29*(2–3), 169–182.

Zhou, Y. B., Li, Q., & Liu, H. Z. (2021). Visual attention and time preference reversals. *Judgment and Decision Making, 16*(4), 1010–1038.